

## CONTENTS

<b>4 Tests of Hypotheses and Confidence Intervals</b>	<b>2</b>
4.1 Test of Hypotheses . . . . .	2
4.2 Hypothesis Tests for Estimable Function . . . . .	5
4.2.1 The Mean Response For Any Treatments . . . . .	6
4.2.2 Difference between the mean response for two treatments	9
4.2.3 Non Estimable Functions .	10
4.3 Consistencies and Redundancies . .	11
4.4 Testable Hypothesis . . . . .	15
4.5 Normal Theory Gauss-Markov Model	17
4.6 Elements of Hypothesis Test . . . .	31
4.6.1 Type I Error Level . . . . .	31
4.6.2 Type II Error Level . . . . .	32
4.6.3 Power of a Test . . . . .	33
4.7 Confidence intervals for estimable functions of $\beta$ . . . . .	39
4.8 Confidence interval for $\sigma^2$ : . . . .	44

## 4 Tests of Hypotheses and Confidence Intervals

### 4.1 Test of Hypotheses

Consider the linear model with

$$E(\mathbf{Y}) = \mathbf{X}\beta \text{ and } Var(\mathbf{Y}) = \Sigma$$

This can also be expressed as

$$\mathbf{Y} = \mathbf{X}\beta + \epsilon$$

where  $E(\epsilon) = \mathbf{0}$  and  $Var(\epsilon) = \Sigma$ .

Typical null hypothesis ( $H_0$ )

- is a status quo or prevailing viewpoint about a population
- specifies the values for one or more elements of  $\beta$
- specifies the values for some linear functions of the elements of  $\beta$

An alternative hypothesis ( $H_1$ )

- is an alternative to the null hypothesis – the change in the population that the researcher hopes is true

We may test

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d} \quad \text{vs} \quad H_1 : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{d}$$

where

- $\mathbf{C}$  is an  $m \times k$  matrix of constants
- $\mathbf{d}$  is an  $m \times 1$  vector of constants

The null hypothesis is rejected if it is shown to be sufficiently incompatible with the observed data.

Failing to reject  $H_0$  is **not** the same as proving  $H_0$  is true.

- too little data to accurately estimate  $\mathbf{C}\boldsymbol{\beta}$
- relatively large variation in  $\boldsymbol{\epsilon}$  (or  $\mathbf{Y}$ )
- if  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$  is false,  $\mathbf{C}\boldsymbol{\beta} - \mathbf{d}$  may be “small”

You can never be completely sure that you made the correct decision

- Type I error (significance level):  
 $P(H_0 \text{ is rejected} | H_0 \text{ is true})$
- Type II error:  
 $P(H_0 \text{ is rejected} | H_0 \text{ is true})$

Basic considerations in specifying a null hypothesis  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$

- $\mathbf{C}\boldsymbol{\beta}$  should be estimable.
- Inconsistencies should be avoided, i.e.,  $\mathbf{C}\boldsymbol{\beta} = \mathbf{d}$  should be a consistent set of equations
- Redundancies should be eliminated, i.e., in  $\mathbf{C}\boldsymbol{\beta} = \mathbf{d}$  we should have

$$\text{rank}(\mathbf{C}) = \text{number-of-rows-in-}\mathbf{C}$$

## 4.2 Hypothesis Tests for Estimable Function

Consider the following effects models:

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij} \quad \begin{matrix} i = 1, 2, 3 \\ j = 1, \dots, n_i \end{matrix}$$

In this case

$$\begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{31} \\ Y_{32} \\ Y_{33} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{21} \\ \epsilon_{31} \\ \epsilon_{32} \\ \epsilon_{33} \end{bmatrix}$$

### 4.2.1 The Mean Response For Any Treatments

By definition

$$E(Y_{ij}) = \mu + \alpha_i \text{ is estimable.}$$

We can test

$$H_0 : \mu + \alpha_1 = 60 \text{ seconds}$$

against

$$H_1 : \mu + \alpha_1 \neq 60 \text{ seconds} \\ \text{(two-sided alternative)}$$

Or we can test

$$H_0 : \mu + \alpha_1 = 60 \text{ seconds}$$

against

$$H_1 : \mu + \alpha_1 < 60 \text{ seconds} \\ \text{(one-sided alternative)}$$

In this case

$$\mu + \alpha_1 = \mathbf{c}^T \boldsymbol{\beta} \quad \text{where} \quad \mathbf{c} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Note that this quantity is estimable, i. e.,

$$\mathbf{c}^T \boldsymbol{\beta} = \mu + \alpha_1 = E \left[ \left( \frac{1}{2} \frac{1}{2} 0 0 0 0 \right) \mathbf{Y} \right].$$

Then, any solution

$$\mathbf{b} = (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \mathbf{Y}$$

to the generalized least squares estimating equations

$$\mathbf{X}^T \Sigma^{-1} \mathbf{X} \mathbf{b} = \mathbf{X}^T \Sigma^{-1} \mathbf{Y}$$

yields the **same value** for  $\mathbf{c}^T \mathbf{b}$  and it is the **unique blue** for  $\mathbf{c}^T \boldsymbol{\beta}$ .

We will reject  $H_0 : \mathbf{c}^T \boldsymbol{\beta} = 60$  if

$$\mathbf{c}^T \mathbf{b} = \mathbf{c}^T (\mathbf{X}^T \Sigma^{-1} \mathbf{X})^{-1} \mathbf{X}^T \Sigma^{-1} \mathbf{Y}$$

is too far away from 60.

If  $\text{Var}(\mathbf{Y}) = \sigma^2 I$ , then any solution

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

to the least squares estimating equations

$$\mathbf{X}^T \mathbf{X} \mathbf{b} = \mathbf{X}^T \mathbf{Y}$$

yields the **same value** for  $\mathbf{c}^T \mathbf{b}$ , and  $\mathbf{c}^T \mathbf{b}$  is the **unique blue** for  $\mathbf{c}^T \boldsymbol{\beta}$ .

We will reject  $H_0 : \mathbf{c}^T \boldsymbol{\beta} = 60$  if

$$\mathbf{c}^T \mathbf{b} = \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

is too far away from 60.

### 4.2.2 Difference between the mean response for two treatments

$$\begin{aligned}\alpha_1 - \alpha_3 &= (\mu + \alpha_1) - (\mu + \alpha_3) \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \frac{-1}{3} & \frac{-1}{3} & \frac{-1}{3} \end{pmatrix} E(\mathbf{Y})\end{aligned}$$

and we can test

$$H_0 : \alpha_1 - \alpha_3 = 0 \quad \text{vs.} \quad H_1 : \alpha_1 - \alpha_3 \neq 0$$

If  $\text{Var}(\mathbf{Y}) = \sigma^2 I$ , the unique BLUE for

$$\alpha_1 - \alpha_3 = (0 \ 1 \ 0 \ -1)\boldsymbol{\beta} = \mathbf{c}^T \boldsymbol{\beta}$$

is

$$\mathbf{c}^T \mathbf{b} \text{ for any } \mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

Reject  $H_0 : \alpha_1 - \alpha_3 = \mathbf{c}^T \boldsymbol{\beta} = 0$  if  $\mathbf{c}^T \mathbf{b}$  is too far from 0.

### 4.2.3 Non Estimable Functions

It would not make much sense to attempt to test

$$H_0 : \alpha_1 = 3 \quad \text{vs.} \quad H_1 : \alpha_1 \neq 3$$

because  $\alpha_1 = [0 \ 1 \ 0 \ 0]\boldsymbol{\beta} = \mathbf{c}^T \boldsymbol{\beta}$  is not estimable.

- Although  $E(Y_{1j}) = \mu + \alpha_1$  neither  $\mu$  nor  $\alpha_1$  has a clear interpretation.
- Different solutions to the normal equations produce different values for

$$\hat{\alpha}_1 = \mathbf{c}^T \mathbf{b} = \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$$

- To make a statement about  $\alpha_1$ , an additional restriction must be imposed on the parameters in the model to give  $\alpha_1$  a precise meaning.

### 4.3 Consistencies and Redundancies

For  $\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$ , consider testing

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \begin{bmatrix} -3 \\ 60 \\ 70 \end{bmatrix} \text{ vs. } H_1 : \mathbf{C}\boldsymbol{\beta} \neq \begin{bmatrix} -3 \\ 60 \\ 70 \end{bmatrix}$$

In this case  $\mathbf{C}\boldsymbol{\beta}$  is estimable, but there is an inconsistency. If the null hypothesis is true,

$$\mathbf{C}\boldsymbol{\beta} = \begin{bmatrix} \alpha_1 - \alpha_3 \\ \mu + \alpha_1 \\ \mu + \alpha_3 \end{bmatrix} = \begin{bmatrix} -3 \\ 60 \\ 70 \end{bmatrix}$$

Then  $\mu + \alpha_1 = 60$  and  $\mu + \alpha_3 = 70$  implies

$$\begin{aligned} (\alpha_1 - \alpha_3) &= (\mu + \alpha_1) - (\mu + \alpha_3) \\ &= 60 - 70 \\ &= \mathbf{-10} \end{aligned}$$

Such inconsistencies should be avoided.

For  $\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

consider testing

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \begin{bmatrix} -10 \\ 60 \\ 70 \end{bmatrix} \text{ vs. } H_1 : \mathbf{C}\boldsymbol{\beta} \neq \begin{bmatrix} -10 \\ 60 \\ 70 \end{bmatrix}$$

In this case  $\mathbf{C}\boldsymbol{\beta}$  is estimable and the equations specified by the null hypothesis are consistent.

There is a redundancy

$$[1 \ 1 \ 0 \ 0] \boldsymbol{\beta} = \mu + \alpha_1 = 60$$

$$[1 \ 0 \ 0 \ 1] \boldsymbol{\beta} = \mu + \alpha_3 = 70$$

imply that

$$\begin{aligned} [0 \ 1 \ 0 \ -1] \boldsymbol{\beta} &= \alpha_1 - \alpha_3 \\ &= (\mu + \alpha_1) - (\mu + \alpha_3) \\ &= 60 - 70 \\ &= -10 \end{aligned}$$

The rows of  $\mathbf{C}$  are not linearly independent, i.e.,  $\text{rank}(\mathbf{C}) < \text{number of rows in } \mathbf{C}$ .

There are many equivalent ways to remove a redundancy:

$$H_0 : \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} 60 \\ 70 \end{bmatrix}$$

$$H_0 : \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 1 & 0 & 0 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} -10 \\ 60 \end{bmatrix}$$

$$H_0 : \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} -10 \\ 70 \end{bmatrix}$$

$$H_0 : \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & 1 & 0 & 1 \end{bmatrix} \boldsymbol{\beta} = \begin{bmatrix} 50 \\ 130 \end{bmatrix}$$

are all equivalent.

In each case:

- The two rows of  $\mathbf{C}$  are linearly independent and

$$\begin{aligned} \text{rank}(\mathbf{C}) &= 2 \\ &= \text{number of rows in } \mathbf{C} \end{aligned}$$

- The two rows of  $\mathbf{C}$  are a basis for the same 2-dimensional subspace of  $R^4$ .

This is the 2-dimensional space spanned by the rows of

$$\mathbf{C} = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

We will only consider null hypotheses of the form

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$$

where  $\text{rank}(\mathbf{C}) = \text{number of rows in } \mathbf{C}$ . This leads to the following concept of a “testable” hypothesis.

## 4.4 Testable Hypothesis

### Definition 1.

Consider a linear model

$$E(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}$$

where

$$V(\mathbf{Y}) = \Sigma$$

and  $\mathbf{X}$  is an  $n \times k$  matrix. For an  $m \times k$  matrix of constants  $\mathbf{C}$  and an  $m \times 1$  vector of constants  $\mathbf{d}$ , we will say that

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$$

is **testable** if

- (i)  $\mathbf{C}\boldsymbol{\beta}$  is estimable
- (ii)  $\text{rank}(\mathbf{C}) = m = \text{number of rows in } \mathbf{C}$

To test  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$

- (i) Use the data to estimate  $\mathbf{C}\boldsymbol{\beta}$ .
- (ii) Reject  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$  if the estimate of  $\mathbf{C}\boldsymbol{\beta}$  is to far away from  $\mathbf{d}$ .
  - How much of the deviation of the estimate of  $\mathbf{C}\boldsymbol{\beta}$  from  $\mathbf{d}$  can be attributed to random errors?
  - Need a probability distribution for the estimate of  $\mathbf{C}\boldsymbol{\beta}$
  - Need a probability distribution for a test statistic



## 4.5 Normal Theory Gauss-Markov Model

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 I)$$

A least squares estimator  $\mathbf{b}$  for  $\boldsymbol{\beta}$  minimizes

$$(\mathbf{Y} - \mathbf{X}\mathbf{b})^T(\mathbf{Y} - \mathbf{X}\mathbf{b})$$

For any generalized inverse of  $\mathbf{X}^T\mathbf{X}$ ,

$$\mathbf{b} = (\mathbf{X}^T\mathbf{X})^- \mathbf{X}^T\mathbf{Y}$$

is a solution to the normal equations

$$(\mathbf{X}^T\mathbf{X})\mathbf{b} = \mathbf{X}^T\mathbf{Y}.$$

## Result 1. Results for the Gauss-Markov model

For a testable null hypothesis

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$$

the OLS estimator for  $\mathbf{C}\boldsymbol{\beta}$ ,

$$\mathbf{Cb} = \mathbf{C}(\mathbf{X}^T\mathbf{X})^- \mathbf{X}^T\mathbf{Y},$$

has the following properties:

- (i) Since  $\mathbf{C}\boldsymbol{\beta}$  is estimable,  $\mathbf{Cb}$  is invariant to the choice of  $(\mathbf{X}^T\mathbf{X})^-$ .
- (ii) Since  $\mathbf{C}\boldsymbol{\beta}$  is estimable,  $\mathbf{Cb}$  is the unique BLUE for  $\mathbf{C}\boldsymbol{\beta}$ .
- (iii)  $E(\mathbf{Cb} - \mathbf{d}) = \mathbf{C}\boldsymbol{\beta} - \mathbf{d}$

(iv)  $V(\mathbf{Cb} - \mathbf{d}) = V(\mathbf{Cb}) = \sigma^2 \mathbf{C}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{C}^T$

(v)  $\mathbf{Cb} - \mathbf{d} \sim N(\mathbf{C}\boldsymbol{\beta} - \mathbf{d}, \sigma^2 \mathbf{C}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{C}^T)$

(vi) When  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$  is true,

$$\mathbf{Cb} - \mathbf{d} \sim N(\mathbf{0}, \sigma^2 \mathbf{C}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{C}^T)$$

(vii) Define

$$SS_{H_0} = (\mathbf{Cb} - \mathbf{d})^T [\mathbf{C}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{C}^T]^{-1} (\mathbf{Cb} - \mathbf{d})$$

then

$$\frac{1}{\sigma^2} SS_{H_0} \sim \chi_m^2(\lambda)$$

where  $m = \text{rank}(\mathbf{C})$  and

$$\lambda = \frac{1}{\sigma^2} (\mathbf{C}\boldsymbol{\beta} - \mathbf{d})^T [\mathbf{C}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{C}^T]^{-1} (\mathbf{C}\boldsymbol{\beta} - \mathbf{d})$$

(viii)  $\frac{1}{\sigma^2} SS_{H_0} \sim \chi_m^2$  if and only if  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$  is true.

(ix)  $E(SS_{residuals}) = (n - k)\sigma^2$  where

$k = \text{rank}(\mathbf{X}) = \text{rank}(\mathbf{P}_\mathbf{X})$  and

$n - k = \text{rank}(\mathbf{I} - \mathbf{P}_\mathbf{X})$  and it follows that

$$MS_{residuals} = \frac{SS_{residuals}}{n - k}$$

is an unbiased estimator of  $\sigma^2$ .

(x)  $\frac{1}{\sigma^2} SS_{residuals} \sim \chi_{n-k}^2$

(xi)  $SS_{H_0}$  and  $SS_{residuals}$  are independently distributed.

$$(xii) \quad F = \frac{\left(\frac{SS_{H_0}}{m\sigma^2}\right)}{\left(\frac{SS_{residuals}}{(n-k)\sigma^2}\right)} = \frac{\frac{SS_{H_0}}{m}}{\frac{SS_{residuals}}{n-k}} = \frac{(n-k)SS_{H_0}}{mSS_{residuals}}$$

$$\sim F_{m,n-k}(\lambda)$$

with noncentrality parameter

$$\lambda = \frac{1}{\sigma^2}(\mathbf{C}\boldsymbol{\beta} - \mathbf{d})^T[\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T]^{-1}(\mathbf{C}\boldsymbol{\beta} - \mathbf{d})$$

$$\geq 0$$

and  $\lambda = 0$  if and only if  $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$  is true.

**Example 1.**

Consider the linear model  $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$  with

$$\mathbf{Y} = \begin{bmatrix} Y_{11} \\ Y_{12} \\ Y_{21} \\ Y_{22} \\ Y_{23} \\ Y_{24} \\ Y_{31} \\ Y_{32} \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

and  $\boldsymbol{\epsilon} \sim N(0, \sigma^2 I)$ .

(a) Determine which of the following hypotheses are testable.

- i.  $H_0 : \alpha_1 = \alpha_2$
- ii.  $H_0 : \alpha_1 - 2\alpha_2 + 3\alpha_3 = 0$
- iii.  $H_0 : \alpha_3 = 0$

iv.  $H_0 : \mu = 0$

v.  $\alpha_1 = \alpha_3$  and  $\alpha_1 - 2\alpha_2 + \alpha_3 = 0$

vi.  $\alpha_1 = \alpha_3$  and  $\alpha_2 = \alpha_3$  and  $\alpha_1 + \alpha_2 - 2\alpha_3 = 0$

(b) Suppose

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \begin{bmatrix} 0 & 1 & 0 & -1 \\ 0 & 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

against the alternative  $H_1 : \alpha_1 \neq \alpha_3$  or  $\alpha_1 - 2\alpha_2 + \alpha_3 \neq 0$ .

i. Show that  $H_0$  is testable.

ii. Express the numerator and denominator of your  $F$ -statistic as two quadratic forms. Show that the quadratic form in the denominator of your  $F$ -statistic, has a central chi-square distribution. Report its degrees of freedom.

- iii. Show that the quadratic form in the numerator of your  $F$ -statistic, has a non central chi-square distribution. Report it's degrees of freedom and non centrality parameter.

- iv. Show that the numerator and denominator of your  $F$ -statistic are independently distributed.

- v. Show that your  $F$ -statistic has a non-central  $F$ -distribution. Report its degrees of freedom and express the non-centrality parameter as a function of  $\alpha_1, \alpha_2, \alpha_3$ .

- vi. Show that your test statistic has a central  $F$ -distribution when the null hypothesis is true.

## 4.6 Elements of Hypothesis Test

We perform the test by rejecting

$$H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{d}$$

if

$$F > F_{(m,n-k),\alpha}$$

where  $\alpha$  is a specified significance level (Type I error level) for the test.

$$\alpha = Pr \{reject H_0 | H_0 \text{ is true}\}$$

### 4.6.1 Type I Error Level

$$\alpha = Pr \{F > F_{m,n-k,\alpha} | H_0 \text{ is true}\}$$

When  $H_0$  is true,

$$F = \frac{MS_{H_0}}{MS_{residuals}} \sim F_{m,n-k}$$

This is the probability of incorrectly rejecting a null hypothesis that is true.

### 4.6.2 Type II Error Level

$$\beta = Pr\{\text{Type II error}\}$$

$$= Pr\{\text{fail to reject } H_0 \mid H_0 \text{ is false}\}$$

$$= Pr\{F < F_{m,n-k,\alpha} \mid H_0 \text{ is false}\}$$

When  $H_0$  is false,

$$F = \frac{MS_{H_0}}{MS_{residuals}} \sim F_{(m,n-k)}(\lambda)$$



### 4.6.3 Power of a Test

$$\begin{aligned} \text{power} &= 1 - \beta \\ &= Pr\{F > F_{m,n-k,\alpha} \mid H_0 \text{ is false}\} \end{aligned}$$

↗  
this determines the value  
of the noncentrality  
parameter.

For a fixed type I error level (significance level)  $\alpha$ , the power of the test increases as the noncentrality parameter increases.

$$\lambda = \frac{1}{\sigma^2}(\mathbf{C}\boldsymbol{\beta} - \mathbf{d})^T[\mathbf{C}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{C}^T]^{-1}(\mathbf{C}\boldsymbol{\beta} - \mathbf{d})$$

### Example 2.

Effects of three diets on blood coagulation times in rats.

Diet factor: Diet 1, Diet 2, Diet 3

Response: blood coagulation time

Model for a completely randomized experiment with  $n_i$  rats assigned to the  $i$ -th diet.

$$Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$$

where

$$\epsilon_{ij} \sim NID(0, \sigma^2)$$

for  $i = 1, 2, 3$  and  $j = 1, 2, \dots, n_i$ .

Here,  $E(Y_{ij}) = \mu + \alpha_i$  is the mean coagulation time for rats fed the  $i$ -th diet.

Test the null hypothesis that the mean blood coagulation time is the same for all three diets.

..

**Example 3.**

Suppose we are willing to specify:

- (i)  $\alpha$  = type I error level = .05
- (ii)  $n_1 = n_2 = n_3 = n$
- (iii) power  $\geq .90$  to detect
- (iv) a specific alternative

$$(\mu + \alpha_1) - (\mu + \alpha_3) = 0.5\sigma$$

$$(\mu + \alpha_2) - (\mu + \alpha_3) = \sigma$$

How many observations (in this case rats) are needed?

**Example 4.**

For the hypotheses testing

$$H_0 : (\mu + \alpha_1) = (\mu + \alpha_2) = \cdots = (\mu + \alpha_k)$$

against

$$H_1 : (\mu + \alpha_1) \neq (\mu + \alpha_j) \text{ for some } i \neq j$$

Obtain the test statistic and the corresponding non-centrality parameter.

## 4.7 Confidence intervals for estimable functions of $\beta$

### Definition 2.

Suppose  $Z \sim N(0, 1)$  is distributed independently of  $W \sim \chi_v^2$ , and then the distribution of

$$t = \frac{Z}{\sqrt{\frac{W}{v}}}$$

is called the student t-distribution with  $v$  degrees of freedom. We will use the notation

$$t \sim t_v$$

For the normal-theory Gauss-Markov model

$$\mathbf{Y} \sim N(\mathbf{X}\beta, \sigma^2 I),$$

the OLS estimator of an estimable function,  $\mathbf{c}^T \beta$ ,

$$\mathbf{c}^T \mathbf{b} = \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-} \mathbf{X}^T \mathbf{Y}$$

follows a normal distribution, i.e.,

$$\mathbf{c}^T \mathbf{b} \sim N(\mathbf{c}^T \beta, \sigma^2 \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-} \mathbf{c}).$$

It follows that

$$Z = \frac{\mathbf{c}^T \mathbf{b} - \mathbf{c}^T \beta}{\sqrt{\sigma^2 \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-} \mathbf{c}}} \sim N(0, 1)$$

From Result 1.(ix), we have

$$\frac{1}{\sigma^2} SSE = \frac{1}{\sigma^2} \mathbf{Y}^T (I - P_{\mathbf{X}}) \mathbf{Y} \sim \chi_{(n-k)}^2$$

where  $k = \text{rank}(\mathbf{X})$ .

Using the same argument that we used to derive Result 1.(x), we can show that  $\mathbf{c}^T \mathbf{b}$  is distributed independently of  $\frac{1}{\sigma^2} SSE$ .

First note that

$$\begin{bmatrix} \mathbf{c}^T \mathbf{b} \\ (I - P_{\mathbf{X}}) \mathbf{Y} \end{bmatrix} \begin{bmatrix} \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-} \mathbf{X}^T \\ (I - P_{\mathbf{X}}) \end{bmatrix} \mathbf{Y}$$

has a joint normal distribution under the normal-theory Gauss-Markov model.

Note that

$$\begin{aligned} & Cov(\mathbf{c}^T \mathbf{b}, (I - P_{\mathbf{X}}) \mathbf{Y}) \\ &= (\mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) (V(\mathbf{Y})) (I - P_{\mathbf{X}})^T \\ &= (\mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T) (\sigma^2) (I - P_{\mathbf{X}}) \\ &= \sigma^2 \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{I} - \mathbf{P}_{\mathbf{X}}) \\ &= 0 \end{aligned}$$

↑  
this is a matrix of zeros

Consequently,

$\mathbf{c}^T \mathbf{b}$  is distributed independently of  
 $\mathbf{e} = (I - P_{\mathbf{X}}) \mathbf{Y}$

which implies that

$\mathbf{c}^T \mathbf{b}$  is distributed independently of  $SSE = \mathbf{e}^T \mathbf{e}$ .

Then,

$$t = \frac{Z}{\sqrt{\frac{SSE}{\sigma^2(n-k)}}}$$

$$\begin{aligned} & \frac{\mathbf{c}^T \mathbf{b} - \mathbf{c}^T \boldsymbol{\beta}}{\sqrt{\sigma^2 \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}}} \\ &= \frac{\mathbf{c}^T \mathbf{b} - \mathbf{c}^T \boldsymbol{\beta}}{\sqrt{\frac{SSE}{\sigma^2(n-k)}}} \\ &= \frac{\mathbf{c}^T \mathbf{b} - \mathbf{c}^T \boldsymbol{\beta}}{\sqrt{\frac{SSE}{(n-k)} \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}}} \sim t_{(n-k)} \end{aligned}$$

↑  
 $\frac{SSE}{n-k}$  is the MSE

It follows that

$$\begin{aligned} 1 - \alpha &= Pr \left\{ -t_{(n-k), \alpha/2} \leq \frac{\mathbf{c}^T \mathbf{b} - \mathbf{c}^T \boldsymbol{\beta}}{\sqrt{MSE \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}}} \leq t_{(n-k), \alpha/2} \right\} \\ &= Pr \left\{ \mathbf{c}^T \mathbf{b} - t_{(n-k), \alpha/2} \sqrt{MSE \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}} \leq \mathbf{c}^T \boldsymbol{\beta} \right. \\ &\quad \left. \leq \mathbf{c}^T \mathbf{b} + t_{(n-k), \alpha/2} \sqrt{MSE \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}} \right\} \end{aligned}$$

and a  $(1 - \alpha) \times 100\%$  confidence interval for  $\mathbf{c}^T \boldsymbol{\beta}$  is

$$\left( \mathbf{c}^T \mathbf{b} - t_{(n-k), \alpha/2} \sqrt{\text{MSE } \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}}, \right. \\ \left. \mathbf{c}^T \mathbf{b} + t_{(n-k), \alpha/2} \sqrt{\text{MSE } \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}} \right)$$

For brevity we will also write

$$\mathbf{c}^T \mathbf{b} \pm t_{(n-k), \alpha/2} S_{\mathbf{c}^T \mathbf{b}}$$

where

$$S_{\mathbf{c}^T \mathbf{b}} = \sqrt{\text{MSE } \mathbf{c}^T (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{c}}.$$

For the normal-theory Gauss-Markov model with  $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 I)$ , the interval

$$\mathbf{c}^T \mathbf{b} \pm t_{(n-k), \alpha/2} S_{\mathbf{c}^T \mathbf{b}}$$

is the **shortest random interval** with probability  $(1 - \alpha)$  of containing  $\mathbf{c}^T \boldsymbol{\beta}$ .

#### 4.8 Confidence interval for $\sigma^2$ :

For the normal-theory Gauss-Markov model with  $\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \sigma^2 I)$  we have shown that

$$\frac{\text{SSE}}{\sigma^2} = \frac{\mathbf{Y}^T (I - P_{\mathbf{X}}) \mathbf{Y}}{\sigma^2} \sim \chi_{(n-k)}^2$$

Then,

$$1 - \alpha = \Pr \left\{ \chi_{(n-k), 1-\alpha/2}^2 \leq \frac{\text{SSE}}{\sigma^2} \leq \chi_{(n-k), \alpha/2}^2 \right\} \\ = \Pr \left\{ \frac{\text{SSE}}{\chi_{(n-k), \alpha/2}^2} \leq \sigma^2 \leq \frac{\text{SSE}}{\chi_{(n-k), 1-\alpha/2}^2} \right\}$$

The resulting  $(1 - \alpha) \times 100\%$  confidence interval for  $\sigma^2$  is

$$\left( \frac{\text{SSE}}{\chi_{(n-k), \alpha/2}^2}, \frac{\text{SSE}}{\chi_{(n-k), 1-\alpha/2}^2} \right)$$

**Example 5.**

For the simple regression model

$$Y_i = \beta_0 + \beta_1 \mathbf{X}_{i1} + \epsilon_i,$$

where for  $\mathbf{e} = (\epsilon_1, \dots, \epsilon_i)^T$ ,  $E(\mathbf{e}) = \mathbf{0}$ . You are given

$$\begin{array}{rcccc} x & 2 & 3 & 4 & 5 \\ y & 4 & 7 & 6 & 8 \end{array}$$

Suppose that  $V(\mathbf{e}) = \sigma^2 I$ .

- (a) Construct the 95% confidence interval for  $\beta_1$ .
- (b) Give 95% two-sided confidence interval for  $\sigma^2$  in the normal version model.